Monte Carlo simulations on the Lefschetz thimble: taming the sign problem

Marco Cristoforetti,[1,2] Francesco Di Renzo,[3] Abhishek Mukherjee,[1,2] and Luigi Scorzato[1,2]

[1] *ECT* / FBK, strada delle tabarelle 286 - 38123 - Trento, Italy.*
[2] *LISC / FBK, via sommarive 18 - 38123 - Trento, Italy.*
[3] *Università di Parma and INFN gruppo collegato di Parma, Viale G.P. Usberti n.7/A - 43124 - Parma, Italy.*

We present the first practical Monte Carlo calculations of the recently proposed Lefschetz thimble formulation of quantum field theories. Our results provide strong evidence that the numerical sign problem that afflicts Monte Carlo calculations of models with complex actions can be softened significantly by changing the domain of integration to the Lefschetz thimble or approximations thereof. We study the interacting complex scalar field theory (relativistic Bose gas) in lattices of size up to 8^4 using a computationally inexpensive approximation of the Lefschetz thimble. Our results are in excellent agreement with known results. We show that—at least in the case of the relativistic Bose gas—the thimble can be systematically approached and the remaining residual phase leads to a much more tractable sign problem (if at all) than the original formulation. This is especially encouraging in view of the wide applicability—in principle—of our method to quantum field theories with a sign problem. We believe that this opens up new possibilities for accurate Monte Carlo calculations in strongly interacting systems of sizes much larger that previously possible.

Introduction — Many important physical systems are characterized by complex actions, when formulated in terms of a path integral. But, if the action S is not real, then e^{-S} is not positive semi-definite and it cannot be interpreted as a probability distribution. In these cases, Monte Carlo calculations are not applicable directly. This is the so called *sign problem.* Many techniques have been proposed to overcome this problem, with important partial successes, but the sign problem is still unsolved for a variety of important physical systems and parameter values, such as lattice QCD at high baryonic density [1], or with a θ-vacuum [2], real-time quantum field theories [3], the electron structure calculations [4–6], the repulsive Hubbard model [7], the nuclear shell model [8] or polymer theory [9], to mention only some of the most famous problems. In this context, any new idea that could improve our chances to simulate any of these models on larger lattices than are feasible today would be extremely valuable.

In a previous work [10, 11], we argued that it may be possible to control the sign problem by reformulating the associated quantum field theory on a Lefschetz thimble. The Lefschetz thimble, associated with a saddle point ϕ, is defined as the hypersurface formed by the union of all paths of steepest descent (SD) of the complex action ending in that saddle point ϕ. Both the Lefschetz thimble and the saddle point are constructed in an enlarged space obtained by complexifying each field component. We showed that, in many cases of interest, this reformulation has the same symmetries and perturbation theory as the original theory [10]. Thereafter, appealing to universality we argued that the reformulation has the same physical content as the original theory.

The benefit of this reformulation is that the action on the Lefschetz thimble has a constant imaginary part, which can be set to zero without any loss of generality. Thus $e^{-\Re\{S\}}$ can now be interpreted as a probability distribution in Monte Carlo sampling. Since, the Lefschetz thimble defines a curved integration domain, there can, in principle, be an additional *residual phase* coming from the Jacobian of the transformation. However, we will argue later that this *residual phase*, if at all present, will result in a very mild growth of stochastic noise.

In this work, we apply our method to the interacting complex scalar field theory describing a relativistic Bose gas at finite chemical potential. This model is one of the simplest non-trivial examples whose sign problem shares many features with the more complex systems mentioned above. Also, in common with lattice QCD, it displays the Silver Blaze phenomenon [12], i.e., the independence of the physics on the chemical potential up to some (finite) critical value. This feature is not accessible to standard Monte Carlo treatments due to the sign problem. Quite importantly, this model has been solved through alternative methods [13–16], and as such provides the ideal test bed for new methods, like ours, for studying the physics of strongly interacting systems.

In the context of Monte Carlo methods, modifications of the domain of integration had been proposed already in [6, 17, 18]. But those deformations were limited to shifts of the contour in the imaginary direction. For many relevant theories, including those considered in [10], the shift is zero, and more general transformations are necessary, to reduce the sign problem. Morse theory [10, 19, 20] identifies the Lefschetz thimbles as the appropriate contours of integration in the more general cases.

Formulation of the model on a Lefschetz thimble — The model is defined by the following continuum action:

$$S = \int d^4x[|\partial\phi|^2 + (m^2 - \mu^2)|\phi|^2 + \mu j_0 + \lambda|\phi|^4], \quad (1)$$

where $\phi(x)$ is a complex scalar field, $j_\nu := \phi^*\partial_\nu\phi - \phi\partial_\nu\phi^*$

and μ is the chemical potential. In this model (as in QCD) the density $\langle n\rangle = \frac{1}{V}\partial \ln Z/\partial\mu$ is expected to be zero up to a critical point. But, this phase transition is hidden in the standard Monte Carlo method because of the strong sign problem which appears as soon as $\mu \neq 0$.

To formulate and simulate the relativistic Bose gas on a Lefschetz thimble [10], we need to discretize the system defined by Eq. (1) and extend the action S holomorphically. This is done by complexifying both the real and imaginary part of the original complex fields $\phi_x = \frac{1}{\sqrt{2}}(\phi_{1,x} + i\phi_{2,x})$, as $\phi_{a,x} = \phi^{(R)}_{a,x} + i\phi^{(I)}_{a,x}$, $a = 1, 2$, which leads to the action in d dimensions [21]:

$$\begin{aligned} S[\{\phi_{a,x}\}] =& \sum_x \Bigg[\left(d + \frac{m^2}{2}\right) \sum_a \phi^2_{a,x} + \frac{\lambda}{4}(\sum_a \phi^2_{a,x})^2 \\ & - \sum_a \sum_{\nu=1}^{d-1} \phi_{a,x}\phi_{a,x+\hat\nu} + \sum_{a,b} i \sinh\mu\, \varepsilon_{ab}\phi_{a,x}\phi_{b,x+\hat0} \\ & - \cosh\mu\, \delta_{a,b}\phi_{a,x}\phi_{b,x+\hat0} \Bigg], \end{aligned} \tag{2}$$

(ε is the 2 dimensional anti-symmetric Levi-Civita symbol). The observables are defined as:

$$\begin{aligned} \langle \mathcal{O}\rangle_0 &= \frac{1}{Z_0}\int_{\mathcal{J}_0} \prod_{a,x} d\phi_{a,x}\; e^{-S[\phi]}\mathcal{O}[\phi], \\ Z_0 &= \int_{\mathcal{J}_0} \prod_{a,x} d\phi_{a,x}\; e^{-S[\phi]}, \end{aligned} \tag{3}$$

where the integration domain $\mathcal{J}_0$ is the Lefschetz thimble [19, 20] attached to $\phi_{\rm glob}$. The configuration $\phi_{\rm glob}$ is the global minimum of the real part of the action $S_R = \Re\{S\}$, when restricted to the original domain $\mathbb{R}^{2V}$. More precisely, $\mathcal{J}_0$ is the manifold of real dimension $N = 2V$, defined as union of all the curves of SD for S_R, i.e., the curves that are solutions of:

$$\begin{aligned} \frac{d}{d\tau}\phi^{(R)}_{a,x}(\tau) &= -\frac{\delta S_R[\phi(\tau)]}{\delta\phi^{(R)}_{a,x}}, \quad \forall a, x, \\ \frac{d}{d\tau}\phi^{(I)}_{a,x}(\tau) &= -\frac{\delta S_R[\phi(\tau)]}{\delta\phi^{(I)}_{a,x}}, \quad \forall a, x, \end{aligned} \tag{4}$$

and that end in $\phi_{\rm glob}$ for $\tau \to \infty$.

In presence of spontaneous symmetry breaking (SSB), the global minimum $\phi_{\rm glob}$ is degenerate. But, the whole procedure can be defined by introducing an explicit term of symmetry breaking: $h\sum_{x,a}\phi_{x,a}$, where h is a real constant, that selects a specific minimum [11] (that can be computed also analytically). Since h is real, the global minimum $\phi_{\rm glob}$ of S_R is also a stationary point of the imaginary part of the action S_I, and hence the thimble is well defined. Physical results are obtained by extrapolating to $h \to 0$.

Aurora Monte Carlo algorithm for sampling the thimble — It is possible to generate field configurations on the Lefschetz thimble with weights given by e^{-S_R} with the help of Langevin dynamics using an algorithm described in [10, 11], that we review here. First, let us assume to know a starting configuration $\phi \in \mathcal{J}_0$, together with a set of configurations $\phi(k\Delta\tau) \in \mathcal{J}_0$, with $k = 1, \ldots, N_\tau$, that represent the path of SD connecting $\phi = \phi(0)$ with the configuration $\phi(\tau = \Delta\tau N_\tau)$. Let us assume that $\phi(\tau)$ is sufficiently close to $\phi_{\rm glob}$, so that the action S can be approximated by its quadratic expansion around $\phi_{\rm glob}$. Second, we generate a Gaussian noise η_j, where $j = 1\ldots 2N$ is a multi-index that stands for $(R/I, a, x)$, and we project it as follows:

$$\eta^\perp_j = P_{j,k}\eta_k, \tag{5}$$

where the $2N \times 2N$ matrix P of rank N is defined in terms of the Hessian matrix H as:

$$P = \frac{H}{\sqrt{H^2}} - 1, \qquad \text{and} \qquad H = \partial^2 S_R[\phi_{\rm glob}]. \tag{6}$$

Then we normalize the noise vector as:

$$\eta' = r\frac{\eta^\perp}{||\eta^\perp||}, \tag{7}$$

where r is a random number distributed according to the N-dimensional χ distribution. This produces a Gaussian noise on the tangent space to $\mathcal{J}_0$ computed in $\phi_{\rm glob}$. We call such linear vector space $\mathcal{G}_0$. Third, we transport the noise from $s = \tau$ along the path of steepest *ascent* (SA) to $s = 0$ by integrating the ordinary differential equation (ODE):

$$\frac{d}{ds}\eta'_j(s) = \sum_k \eta'_k(s)A(s)_{k,j}, \tag{8}$$

where $A(s)_{k,j}$ is the Iwasawa projection of $\partial_k\partial_j S_R[\phi(s)]$ [11]. This ensures that the noise remains tangent to $\mathcal{J}_0$. Fourth, we use the evolved noise $\eta'(0)$ to generate a new configuration as:

$$\phi'_j = \phi_j - \Delta t\,\frac{\delta S_R[\phi]}{\delta\phi_j} + \sqrt{2\Delta t}\,\eta'_j.$$

In the limit $\Delta t \to 0$ this simulates Langevin dynamics on the thimble. For $\Delta t > 0$, $\phi'(0)$ will move away from the thimble of order $(\Delta t)^2$. To correct this, the fifth step consists in following the path of SD from $\phi'(0)$ for a length τ leading to the configuration $\phi'(\tau)$. Assuming that the action at $\phi'(\tau)$ can be approximated with its quadratic part (otherwise, we extend τ), we ensure that $\phi'(\tau)$ belongs to the thimble by projecting it as $\phi(\tau)^{(\rm new)} = P\phi'(\tau)$. Finally, we follow the path of SA from $\phi(\tau)^{(\rm new)}$ for a length τ. The resulting $\phi(0)^{(\rm new)}$ is the new configuration[1].

[1] Note that this procedure is not inherently stable, as the one in [10], but relies on the (verifiable) fact that the integration in τ always brings sufficiently close to the saddle point.

The computation of the projector P is done, once for all, at the beginning of the simulation. However, it must be applied at every iteration. This can be done most efficiently in Fourier space, where H and P are diagonal, although, for this first exploratory study on small lattices, we did not take advantage of this possibility.

The cost of the algorithm depends significantly on the length τ, that should be large enough to stretch out to the region where the quadratic approximation of the action is good. But how good is good enough? We certainly do not need to constrain the system on the thimble exactly, but only to the extent that the domain of integration preserves the homology class of the thimble and the reweighting with the phase e^{iS_I} is feasible.

It is then natural to ask whether $\tau = 0$ is already sufficient. This corresponds to integrating the system on the vector space $\mathcal{G}_0$ defined above. In general, $\mathcal{G}_0$ does not belong to the same homology class as $\mathcal{J}_0$, because the directions of steepest ascent for the quadratic part of the action may not, in general, be directions of convergence for the full action.

However, in our simulations we observed that such divergences, although they do occur as expected, are very rare (see below). This suggests that the integration on $\mathcal{G}_0$, regularized, say, with a mild cutoff, might already provide a good approximation. Of course, such a regulator introduces an unknown bias, and the procedure is meaningful only if the regulator is eventually removed, by approaching the true thimble further. Next, we present our results on $\mathcal{G}_0$, following which we show how the true thimble can be systematically approached.

Numerical results on $\mathcal{G}_0$ — As discussed above, the simulations on $\mathcal{G}_0$ are meaningful only with a regulator. Instead of introducing an explicit cut of the domain, we regularized by discarding those simulations that diverged within the observed histories (i.e. 4×10^6 trajectories for $V = 4^4$, 10^6 trajectories for $V = 6^4$ and 8×10^5 trajectories for $V = 8^4$). This procedure introduces an unknown bias, that can only be removed by approaching the thimble further. However, the fact that the divergences are very rare makes the regularization rather unambiguous. If we consider a common span of the first 8×10^5 trajectories, a divergence occurred with probability $\sim 1.8\%$ on the lattices $V = 4^4$, with probability $\sim 0.8\%$ on $V = 6^4$, and less than 0.7% on $V = 8^4$ ($h = 5 \times 10^{-3}$). The results obtained in this way agree perfectly (within the rather small errors) with the results obtained with the algorithm of [15] and [14][2]. In particular, they show the correct scaling with the volume. Note that, since $\mathcal{G}_0$ is a flat manifold, the *residual phase*, discussed in [10, 11] is absent.

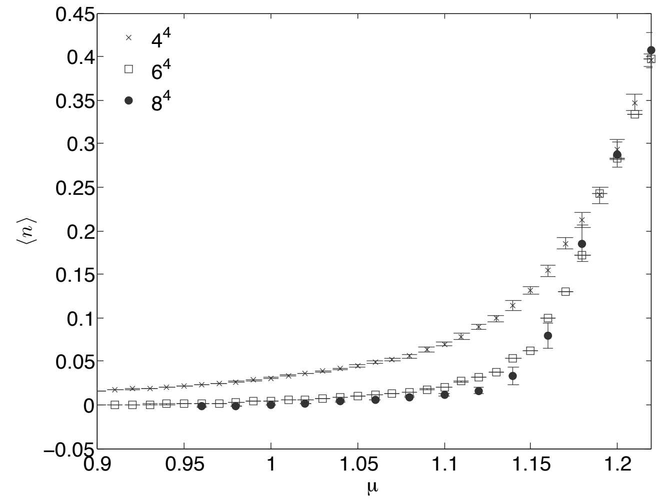

FIG. 1. Average density $\langle n \rangle$ in the critical region for the lattices $V = 4^4, 6^4, 8^4$.

We report the results of simulations for the relativistic Bose gas in $3 + 1$ dimensions ($d = 4$). The mass and coupling were fixed at $m = 1 = \lambda$, and μ was varied from 0 to 1.3. In Fig. 1 and 2, we plot our results for the density $\langle n \rangle$ and $\langle |\phi|^2 \rangle$ in the most interesting range between $\mu = 0.9$ and $\mu = 1.22$. In these figures, we see a clear signal of the Silver Blaze phenomenon around $\mu \sim 1.1$. In all the simulations shown here we used $\Delta t = 10^{-4}$, but we performed also some tests with $\Delta t = 10^{-3}$ and $\Delta t = 10^{-5}$ and we found no significant difference. The errorbars on each point are computed from the standard deviation of 10 to 20 independent histories, in order to take the autocorrelation effects into account. We used the sources $h = 5 \times 10^{-3}$ and $h = 10^{-3}$ to extract the limit $h \to 0$.

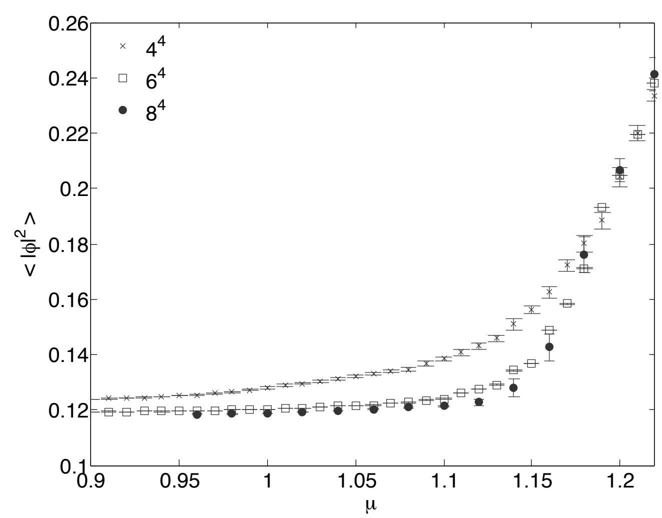

FIG. 2. Same as in Fig. 1 for the observable $\langle |\phi|^2 \rangle$.

In Fig. 3 we plot the average phase for the same sim-

[2] We thank Gert Aarts, Christof Gattringer and Thomas Kloiber for sharing their partially unpublished results with us.

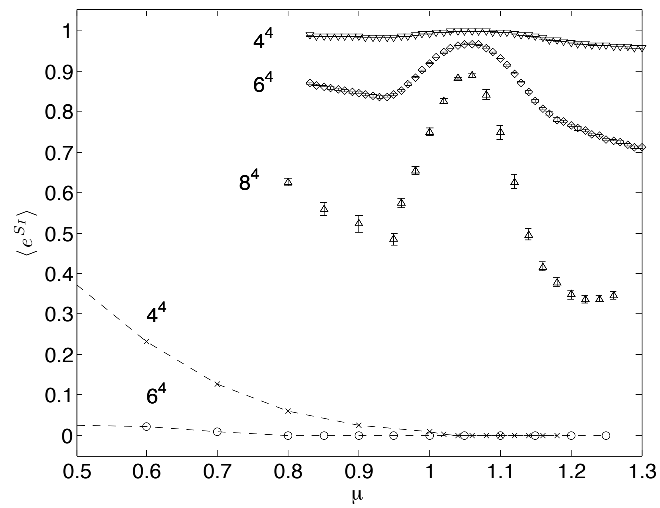

FIG. 3. The data on the top-right show the average phase obtained with the Aurora algorithm on lattices 4^4, 6^4 and 8^4. It is interesting that the average phase is large precisely in the most interesting region just above $\mu = 1$. The dashed lines on the bottom-left display, for comparison, the average phase obtained with a naive phase-quenched Monte Carlo algorithm on lattices 4^4 and 6^4. Even on a 4^4 lattice, the sign problem in the phase-quenched algorithm, completely hides the interesting region.

ulations reported above. The phase is used to reweight the observables. However, such reweighting brings corrections to the observables that are unnoticeable, within the statistical errors. As expected, the sign problem in $\mathcal{G}_0$ gradually increases on larger volumes and moving closer to the thimble will be eventually necessary.

Moving closer to the thimble — In general, there are two good reasons to move closer to the thimble $\mathcal{J}_0$. First, to remove the bias introduced by the regulator on $\mathcal{G}_0$. Second, to keep the sign problem under control on larger volumes. However, in the present situation, the divergences are already very rare and to observe a further measurable reduction would require enormous statistics. Moreover, the results obtained on $\mathcal{G}_0$ are already in excellent agreement with the known results, and the reweighting with the phase has no effect even in the most critical case of the 8^4 lattice at $\mu = 1.2$. Hence, the results reported here with $\tau \neq 0$ do not intend to improve the precision of the results obtained above with $\tau = 0$, but rather to present a first exploration of the feasibility of moving closer to the thimble.

To integrate Eqs. (4) and (8), we employed the (classical) 4th order Runge-Kutta method (RK4). This is an explicit method, that can be used to solve Eq. (4) and (8) as initial value problems (IVP). We argued in [10] that, in order to enable a stable integration in the most general case for large τ, without the need of too tiny $\Delta\tau$, Eq. (4) should be treated as a boundary value problem (BVP), by introducing explicit boundary conditions in the neighborhood of the saddle point. However, it is interesting to see what can be achieved even with the simpler procedure adopted here.

To evaluate the closeness to the thimble, we monitored the reduction of the fluctuations of the imaginary part of the action (what really matters for the sign problem), when τ is increased. We found that $\tau = 4 \times 10^{-2}$ was sufficient to suppress the fluctuations of the imaginary part of the action S_I by a factor ~ 0.5 (for 4^4), a factor ~ 0.6 (for 6^4), and ~ 0.7 (for 8^4). This test was performed for $\mu = 1.2$, in the critical region. However, a precise integration of the IVP becomes more and more difficult on increasingly large volumes (the correctness of the integrator can be assessed via reversibility checks). This shows that the IVP formulation of the ODE will need to be replaced by a BVP formulation in more difficult situations.

Finally, note that in this test we neglected the computation of the *residual phase* discussed in [10, 11]. But the excellent agreement with the known results, even without including the residual phase, supports the idea that its effect is not dramatic and maybe even negligible.

Summary — We have reported the first numerical application of the Lefschetz formulation to a nontrivial model with a hard sign problem. In particular, we have studied the relativistic Bose gas model at finite chemical potential. Our study was restricted to small lattices, but, given the severity of the sign problem, this can be considered already a very challenging test. We found excellent agreement with the known results already on the crudest approximation of the thimble, i.e, the vector space $\mathcal{G}_0$, once the integral was regulated by removing the few diverging trajectories. Moreover, we showed that it is possible to improve the approximation of the thimble, by following the equations of SD.

Of course, the sign problem is expected to become worse on larger lattices: moving closer to the thimble will become more crucial. Work in progress include developing efficient and stable integration algorithms to achieve a better approximation of the thimble, study of the scaling for larger system sizes, and application of our method to other models.

Acknowledgments — We are grateful to Christian Torrero for many valuable discussions and to Gert Aarts, Christof Gattringer and Thomas Kloiber for communications about their results. This research is partially supported by the PAT-INFN project AuroraScience, by the EU under Grant Agreement No. PITN-GA-2009-238353 (ITN STRONGnet), by INFN through the i.s. MI11 and by MIUR under contract PRIN2009 (20093BM-NPR_004). Numerical simulations were performed with the Aurora prototype at FBK.

[1] G. Aarts, PoS **LATTICE2012**, 017 (2012), arXiv:1302.3028 [hep-lat].

[2] E. Vicari and H. Panagopoulos, Phys.Rept. **470**, 93 (2009), arXiv:0803.1593 [hep-th].

[3] J. Berges, S. Borsanyi, D. Sexty, and I.-O. Stamatescu, Phys.Rev. **D75**, 045007 (2007), arXiv:hep-lat/0609058 [hep-lat].

[4] P. L. Silvestrelli, S. Baroni, and R. Car, Phys. Rev. Lett. **71**, 1148 (1993).

[5] D. M. Charutz and D. Neuhauser, J. Chem. Phys. **102**, 4495 (1995).

[6] R. Baer, M. Head-Gordon, and D. Neuhauser, J. Chem. Phys. **109**, 6219 (1998).

[7] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[8] S. Koonin, D. Dean, and K. Langanke, Phys. Rept. **278**, 1 (1997).

[9] T. Vilgis, Phys. Rept. **336**, 167 (2000).

[10] M. Cristoforetti, F. Di Renzo, and L. Scorzato (Aurora-Science Collaboration), Phys. Rev. D **86**, 074506 (2012), arXiv:1205.3996 [hep-lat].

[11] M. Cristoforetti, F. Di Renzo, and L. Scorzato, (2012), arXiv:1210.8026 [hep-lat].

[12] T. D. Cohen, Phys.Rev.Lett. **91**, 222001 (2003), hep-ph/0307089.

[13] M. G. Endres, Phys.Rev. **D75**, 065012 (2007), hep-lat/0610029.

[14] G. Aarts, Phys.Rev.Lett. **102**, 131601 (2009), arXiv:0810.2089 [hep-lat].

[15] C. Gattringer and T. Kloiber, Nucl.Phys. **B869**, 56 (2013), arXiv:1206.2954 [hep-lat].

[16] C. Gattringer and T. Kloiber, (2012), arXiv:1212.3770 [hep-lat].

[17] N. Rom, E. Fattal, A. K. Gupta, E. A. Carter, and D. Neuhauser, J. Chem. Phys. **109**, 8241 (1998).

[18] A. G. Moreira, S. A. Baeurle, and G. H. Fredrickson, Phys. Rev. Lett. **91**, 150201 (2003).

[19] F. Pham, Proc. Symp. Pure Math. **40**, 319 (1983).

[20] E. Witten, (2010), arXiv:1001.2933 [hep-th].

[21] G. Aarts, JHEP **0905**, 052 (2009), arXiv:0902.4686 [hep-lat].